UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF TEXAS
FORT WORTH DIVISION

ELEVANCE HEALTH, INC., ET AL.,

Plaintiffs,

v. **No. 4:24-cv-01064-P**

ROBERT F. KENNEDY, JR., IN HIS OFFICIAL CAPACITY AS SECRETARY OF HEALTH AND HUMAN SERVICES, ET AL.,

Defendants.

OPINION & ORDER

Before the Court are two cross-motions for summary judgment—one filed by Plaintiffs Elevance Health, Inc.; Community Insurance Company; Wellpoint Insurance Company; Wellpoint Texas, Inc.; Freedom Health, Inc.; and Group Retiree Health Solutions (collectively, Elevance); and the other filed by Defendants Robert F. Kennedy, Jr., in his capacity as Secretary of the United States Department of Health and Human Services, and Stephanie Carlton in her capacity as Acting Administrator of the Centers for Medicare and Medicaid Services (collectively, CMS). ECF Nos. 31, 37. After considering the motions, the briefs, the administrative record, and the applicable law, the Court will **DENY** Elevance's motion and **GRANT** CMS's motion.

BACKGROUND

CMS assigns a Star Rating to every contract it makes with a Medicare Advantage Organization (MAO). CMS issues the Star Ratings in half-star increments based on statutory and regulatory guidelines. For MAOs like Elevance, a mere half-star difference can mean the difference in hundreds of millions of dollars in potential federal funding. Elevance contends that CMS arbitrarily and capriciously kept some of its contracts from reaching the next highest half-star tier in the 2025 Star Ratings. To provide a clear understanding of Elevance's claims, the Court starts with the statutory and regulatory background.

A. Statutory and Regulatory Background

In 1965, Congress amended the Social Security Act to create the Medicare Program, through which the elderly and disabled can receive health insurance from the federal government. 42 U.S.C. §§ 1395 *et seq*. The Center for Medicare and Medicaid Services, a branch of the United States Department of Health and Human Services, administers the program. Medicare is made up of four parts. Part C, the one relevant here, is the Medicare Advantage program. *Id.* §§ 1395w-21–1395w-29.

Medicare Advantage lets patients choose to receive Medicare benefits from private health insurance providers known as Medicare Advantage Organizations. Each MAO contracts with CMS to provide coverage to enrollees in a given geographic area. MAOs make bids to CMS to provide coverage for a lower price than CMS's benchmark rate. *Id.* § 1395w-23(n); 42 C.F.R. § 422.258. In exchange for providing cheaper coverage, MAOs receive a portion of the difference between their bids and the benchmark. 42 U.S.C. § 1395w-23(a)(1)(B)(i). The size of that portion depends, in part, on each contract's Star Rating. 42 U.S.C. §§ 1395w-23(o), 1395w-24(b)(1)(C)(v).

1. Star Ratings

CMS assigns a Star Rating to each MAO contract based on a host of measurements related to quality of coverage, ease of access, beneficiary experience, and other aspects. 42 U.S.C. § 1395w-23(o)(4)(A). The Star Rating system "is designed to provide information to the beneficiary that is a true reflection of the plan's quality and encompasses multiple dimensions of high-quality care." 83 Fed. Reg. at 16,520. Medicare Advantage patients can look at the Star Ratings of competing providers to compare and choose between them. The Star Rating of a MAO's contract also determines the amount of the direct payment to the MAO as well as the percentage of the savings the MAO will receive in the form of a rebate. 42 U.S.C. §§ 1395w-23(o), 1395w-24(b)(1)(C)(v). Every MAO contract's rating is measured in half-star increments (1 star, 1.5 stars, and so on). 42 C.F.R. § 422.166(c)(3). The process used to assign Star Ratings to MAO contracts is complex. The following background section, while detailed, is still only a summary. But the Court must explain the regulatory background to intelligibly explain the disputes in this case.

Each contract's overall Star Rating is a complex composite of several *measure-level ratings*. CMS assigns measure-level ratings to each Part C contract based on around 30 quality measures. Each quality measure is given a measure-level rating, which is recorded in whole-star increments. 42 C.F.R. § 422.166(a)(4).

The measure-level ratings quantify many characteristics of MAO contracts, each of which is weighted by its relative importance. Improvement measures—those based on the changes of a contract's measure scores from year to year—receive the greatest weight. Other kinds of ratings measure patient experience, complaints, ease of access, patient outcomes, and processes. CMS takes all of a contract's measure-level ratings, assigns each one a certain weight, and takes the average of those values. That average is the raw score that gets rounded to form a summary Star Rating (and, if the contract is rated for both Parts C and D, an overall Star Rating) for that contract.

CMS gathers the data needed for the Star Ratings in several ways. Broadly, the measures fall into two boxes: survey-based data gathered from the Consumer Assessment of Healthcare Providers and Systems (CAHPS), and other administrative and medical record review data, including data collected as part of the Healthcare Effectiveness Data and Information Set (HEDIS). These two sources are called *CAHPS measures* and *non-CAHPS measures*, respectively. CMS uses different processes to convert those sources of data into measure-level ratings.

2. CAHPS Measures

For CAHPS measures, CMS applies a *case-mix adjustment*. Case-mix adjustment is "an adjustment to the measure score made prior to the score being converted into a Star Rating to take into account certain enrollee characteristics that are not under the control of the plan." 42 C.F.R. § 162(a). Case-mix adjustment aims to control for traits of a contract's enrollees that would otherwise tend to skew the data from that pool. (For an unsurprising example, enrollees who report better health overall tend to provide more positive reports of the healthcare they receive. ECF No. 33 at App. 189.) Case-mix adjustment accounts for those kinds of trends by bumping each measure up or down based on the demographics of the contract's enrollees.

Next, CMS turns case-mix adjusted CAHPS data into measure-level ratings by a method called *relative distribution and significance testing* (RDS Testing). 42 C.F.R. §§ 422.166(a)(3), 423.186(a)(3). For every measure, CMS first determines how reliable the survey data is, the percentile into which that measure falls compared to all other contracts, and its distance from the national average. A formula, codified at 42 C.F.R. § 422.166(a)(3), then combines those three values (survey reliability, percentile, and number of standard errors from the national average) to determine that measure's whole-number rating.

3. Non-CAHPS Measures

CMS processes non-CAHPS data differently. CMS gathers data for non-CAHPS measures from HEDIS as well as from administrative data, data collected from contractors, and other surveys. Instead of using RDS Testing, CMS uses a clustering algorithm to turn non-CAHPS measures into measure-level ratings. The algorithm's goal is to produce measure-level ratings with as little variation within each star—and as much variation from one star to the next—as statistically possible. It works by identifying the most statistically significant gaps among the scores and creates four *cut points*, resulting in five groups—one for each measure-level rating. In principle, the algorithm avoids a result in which the cutoff between one star rating and the next is in the middle of a tight cluster of scores. Instead, the cutoff between one star and the next would ideally land in a relatively wide gap.

Regulations require the non-CAHPS cut points to be refined using a process called *mean resampling*. Mean resampling ensures that the cut points do not change drastically from one year to the next or fluctuate too much based on individual contract scores. In this method, all the scores for a given non-CAHPS measure are randomly reshuffled into ten equally sized groups. The algorithm is run ten times, leaving one of the ten groups out each time. Ten sets of four cut points are produced. The ten sets are then averaged to produce four refined cut points. After the mean resampling process is complete, the cut points are determined, and the non-CAHPS measures are sorted into five tiers and assigned measure-level ratings accordingly.

The random assortment of contracts into ten groups depends on a *seed* number. The seed is a sequence of numbers that kickstarts the randomization process by sorting contracts into ten groups. CMS uses the same seed every year for mean resampling. The number it uses is 8,675,309, an allusion to the popular song "867-5309 (Jenny)" by musical artist Tommy Tutone. (For that reason, it is sometimes called the "Jenny Seed.")

4. Final Star Ratings

After the measure-level ratings for contracts have been calculated, CMS weights the ratings by importance and then takes the average. The result is a raw score between one and five stars. Because it is the average of several weighted measurements, the exact number can have (and presumably often does have) a long or non-terminating decimal. CMS calculates the raw score to at least six decimal places. CMS therefore rounds a contract's raw score to the nearest half star, as required by statute. Regulations specify that CMS must follow "traditional rounding rules." 42 C.F.R. §§ 422.166(d)(2)(iv); 422.162(a). The parties do not agree on how the traditional rules apply when rounding to the nearest half.

B. Factual History

Elevance is a healthcare company based in Indianapolis. Through several subsidiaries (including some Plaintiffs here), Elevance operates health plans, including Medicare Advantage Plans, around the country. Elevance is the "parent organization" for five contracts with CMS at issue in this case: H3655, H5427, H6078, H2593, and H8849. *See* 42 C.F.R. § 422.2. When CMS released the Star Ratings for 2025, the five contracts at issue were rated as follows (ECF No. 33 at App. 11):

H2593	3.5 Stars
H3655	3.5 Stars
H5427	4 Stars
H6078	3 Stars
H8849	3.5 Stars

Elevance sued HHS and CMS, arguing that the calculations of the 2025 Star Ratings were arbitrary and capricious and contrary to law, in violation of the Administrative Procedure Act. ECF No. 22. Specifically, Elevance argues as follows: *First*, CMS deviated from its statutory and regulatory authority when it adjusted the CAHPS measure scores. *Second*, CMS did not follow traditional rounding rules correctly when it rounded the Star Ratings, causing contract H3655 to be rounded down to 3.5 Stars instead of 4 Stars. Under those two headings, Elevance makes multiple arguments based on several actions taken by CMS during the preparation of the Star Ratings. Both Elevance and CMS now move for summary judgment. The motions are ripe for review.

LEGAL STANDARD

The Administrative Procedure Act (APA) provides that, "[t]o the extent necessary to decision and when presented, the reviewing court shall decide all relevant questions of law, interpret constitutional and statutory provisions, and determine the meaning or applicability of the terms of an agency action." 5 U.S.C. § 706. In a case challenging an agency action under the APA, summary judgment "serves as the mechanism for deciding" whether the action "is supported by the administrative record and otherwise consistent with the APA standard of review." *Gadhave v. Thompson*, No. 3:21-cv-2938-D, 2023 WL 6931334, at *1 (N.D. Tex. Oct. 19, 2023) (citation omitted). The United States Court of Appeals for the Fifth Circuit has "consistently upheld, without comment, the use of summary judgment as a mechanism for review of agency decisions." *Girling Health Care, Inc. v. Shalala*, 85 F.3d 211, 214 (5th Cir. 1996).

The agency resolves "factual issues to arrive at a decision supported by the administrative record." *Yogi Metals Grp. Inc. v. Garland*, 567 F. Supp. 3d 793, 797–98 (S.D. Tex. 2021), *aff'd,* 38 F.4th 455 (5th Cir. 2022) (citation omitted). The district court then applies the APA standards of review to determine whether, as a matter of law, "the evidence in the administrative record permitted the agency to make the decision it did." *MRC Energy Co. v. U.S. Citizenship & Immigr. Servs.*, No. 3:19-cv-2003-K, 2021 WL 1209188, at *3 (N.D. Tex. Mar. 31, 2021) (citation omitted). "The entire case is thus a question of law, with the district court sitting

as an appellate tribunal." *Outsourcing Facilities Ass'n v. U.S. Food & Drug Admin.*, No. 4:24-cv-0953-P, 2025 WL 1397537 at *1 (N.D. Tex. May 13, 2025) (Pittman, J.).

Agency decisions are "presumptively valid; the [plaintiff] bears the burden of showing otherwise." *Barr v. SEC*, 114 F.4th 441, 447 (5th Cir. 2024); *Tex. Med. Ass'n v. U.S. Dep't of Health & Hum. Servs.*, 120 F.4th 494, 504 (5th Cir. 2024) (internal quotation and citation omitted). "If the agency articulates a rational relationship between the facts found and the choice made it does not act arbitrarily or capriciously." *Joseph v. Dir. of Tex. Serv. Ctr., U.S. Citizenship & Immigr. Servs.*, No. 24-40249, 2025 WL 458001, at *3 (5th Cir. Feb. 11, 2025) (quoting *Louisiana ex rel. Guste v. Verity*, 853 F.2d 322, 327 (5th Cir. 1988)). The "focal point" of that review "should be the administrative record already in existence, not some new record made initially in the reviewing court." *Camp v. Pitts*, 411 U.S. 138, 142 (1973). And "[j]udicial review under that standard is deferential, a[s] a court may not substitute its own policy judgment for that of the agency." *FCC v. Prometheus Radio Project*, 592 U.S. 414, 423 (2021). Although courts "may not supply a reasoned basis for the agency's action that the agency itself has not given," courts are to "uphold a decision of less than ideal clarity if the agency's path may reasonably be discerned." *Tex. Med. Ass'n,* 120 F.4th at 504 (citing *Motor Vehicle Mfrs. Ass'n of U.S. v. State Farm Mut. Auto. Ins.*, 463 U.S. 29, 43 (1983)).

Still, a review of an APA challenge has "serious bite." *Texas v. Env't Prot. Agency*, 91 F.4th 280, 291 (5th Cir. 2024). That holds true especially when an agency is alleged to have acted contrary to law. "The failure of an agency to follow its regulations renders its decision invalid." *Gulf States Mfrs., Inc. v. Nat'l Lab. Rels. Bd.*, 579 F.2d 1298, 1308 (5th Cir. 1978). The Court does not yield to an agency's interpretation of its own regulations, even if the regulation is ambiguous. *Loper Bright Enterps. v. Raimondo*, 603 U.S. 639, 412–13 (2024). Instead, the Court interprets the agency's regulation using its "independent judgment." *Id.* at 413. In short, the Court reviews agency actions to ensure they comply with the law and do not blatantly disregard the administrative record. The Court will do so below.

ANALYSIS

At summary judgment, Elevance makes two main contentions: CMS erroneously adjusted the measure-level ratings for CAHPS measure scores and improperly rounded each contract's raw score to the sixth decimal place before assigning it a Star Rating. Under the first heading, there are two main complaints. *First*, CMS applied case-mix adjustment to CAHPS data, which it was not authorized to do. And *second*, CMS compared each contract's CAHPS measure data to the national weighted average, as opposed to the contract-level average. Under the next heading, Elevance brings three more arguments. *First*, CMS should have rounded the score of contract H3655 from 3.749565 to 4, not 3.5. *Second*, rounding to the sixth decimal place was arbitrary given the imprecision of the scoring calculations. *Third*, rounding to the sixth decimal place was arbitrary given the use of the Jenny Seed year after year. The Court will take all five of those arguments in turn.

A. Case-Mix Adjustment

Case-mix adjustment is the process whereby CMS adjusts a certain CAHPS measure score before converting it to a measure-level rating. The results of a survey can be skewed by the makeup of the respondents. Because some characteristics of respondents might affect their survey answers for reasons unrelated to the quality of their care, CMS adjusts each CAHPS measure score. Case-mix adjustments account for characteristics such as age, education, chronic medical conditions, and functional health status. The process aims to remove bias from CAHPS measure scores.

Elevance does not argue that the case-mix adjustment process, in and of itself, is arbitrary or capricious. Instead, it argues that the case-mix adjustment of CAHPS measures is not authorized by any statute or regulation. Other regulations, Elevance argues, explicitly authorize case-mix adjustments for different measures. But case-mix adjustments are not expressly authorized for CAHPS measures. Therefore, Elevance contends, CMS deviated from the regulations when it adjusted the CAHPS scores for case-mix indexes. And if CMS deviated from its enabling statutes or regulations, it per se acted arbitrarily and capriciously. *See Gulf States Mfrs.*, 579 F.2d at 1308.

Elevance further contends it was harmed by the use of case-mix adjustments. Elevance submits the testimony of J. Mark Abernathy, managing director of Berkeley Research Group, whom it proposes as an "expert in the managed care industry." ECF No. 33 at App. 1–2. Abernathy testified that he "recalculate[d] the CAHPS measures for the 2025 Star Ratings without the use of the case-mix adjustment . . . ." *Id.* at App. 2. Abernathy concluded that if CMS had not made case-mix adjustments to the CAHPS measures, two of Elevance's contracts—H3655 and H6078—would have received a score one half-star higher. *Id.* at App. 16.

In response, CMS overstates its case that the regulations clearly authorize case-mix adjustment. CMS points out that the regulations define "case-mix adjustment" using the past participle "made" ("an adjustment to the measure score made prior to the score being converted"). It argues that use of the past participle "made" means that it has the authority to make case-mix adjustments at its discretion. That does not follow.[1]

In fact, the regulations do not specify which measures should be case-mix adjusted. But the regulation governing Star Ratings at several points assumes that various measures will have been adjusted already. For example, one section discusses the categorical adjustment index (CAI), a different adjustment not otherwise relevant here. Excluded from the CAI is any measure that "is already case-mix adjusted for socioeconomic status." *Id.* § 422.166(f)(2)(ii)(A). But case-mix adjustments based on socioeconomic status are not expressly authorized anywhere. Likewise, another section describes the health equity index, another adjustment based on social risk factors. *Id.* § 422.166(f)(3). That section also refers to measures "that are case-mix adjusted in the Star Ratings" and provides special rules for those measures when applying the health equity index adjustment. *Id.* § 422.166(f)(3)(i)(A).

[1]CMS's argument commits the *existential fallacy*, in which one presupposes that a class has members without warrant to do so. *See* "The Existential Fallacy," FALLACY FILES, https://www.fallacyfiles.org/existent.html, last visited Aug. 7, 2025.

This regulatory background undercuts Elevance's argument. Elevance argues that because CMS referred to case-mix adjustment in certain parts of the regulations but not for CAHPS measure scores, the implication is that adjustment is not authorized for CAHPS measure scores. But the text of the regulations cuts the other way: the few references to case-mix adjustments in the regulations assume it is already taking place. Those references imply that case-mix adjustments belong to CMS's discretion to add and update measures through notice-and-comment rulemaking.[2] *See* 42 C.F.R. § 422.164.

Although the regulations do not expressly authorize case-mix adjustments for CAHPS measure scores, they refer to case-mix adjustment in a way that assumes it is taking place already. Besides, it may be arbitrary or capricious for CMS *not* to make case-mix adjustments to survey data. As Elevance repeats often, an agency acts arbitrarily and capriciously if it "entirely fail[s] to consider an important aspect of [a] problem . . . ." *Motor Vehicle Mfrs. Ass'n*, 463 U.S. at 43. Arguably, the variation between the demographics of different contracts is an important aspect of a problem that would be arbitrary for CMS to ignore. Tellingly, Elevance never argues that making case-mix adjustments is per se arbitrary or capricious; it is left only with its argument from statutory and regulatory silence. And that argument fails because CMS's regulations assume measures are case-mix adjusted—albeit without expressly authorizing them.

CMS acted within its statutory authority when it applied a case-mix adjustment to the CAHPS measure scores for the 2025 Star Ratings. It is therefore entitled to summary judgment on Elevance's arbitrary-and-capricious claim on this ground.

---

[2]Additionally, the regulations provide that CMS will "list[ ] the measures used for a particular Star Rating each year in the [Medicare Part C & D Star Ratings] Technical Notes or similar guidance document with publication of the Star Ratings." *Id.* § 422.164(a). CMS identifies whether each measure is case-mix adjusted in their annual Technical Notes. *See* Administrative Record (A.R.) at 38–113.

B. How to Calculate the National Average

Elevance complains that in addition to improperly making case-mix adjustments to CAHPS measure scores, CMS effectively penalized each contract again by using a national *weighted* average when performing RDS Testing. As discussed above, RDS Testing looks at how a contract's average CAHPS measure score relates to the "national average." The regulations provide a formula for turning a case-mix adjusted CAHPS measure score into a whole-number measure-level rating. For example:

> (i) A contract is assigned 1 star if both of the criteria in paragraphs (a)(3)(i)(A) and (B) of this section are met plus at least one of the criteria in paragraphs (a)(3)(i)(C) or (D) of this section is met:
>
> > (A) Its average CAHPS measure score is lower than the 15th percentile; and
> >
> > (B) Its average CAHPS measure score is statistically significantly lower than the national average CAHPS measure score;
> >
> > (C) The reliability is not low; or
> >
> > (D) Its average CAHPS measure score is more than one standard error below the 15th percentile.

42 C.F.R. § 422.166(a)(3)(i). In other words, a contract will receive one star for a measure if three conditions are met. *First*, it falls below the 15th percentile. *Second*, it is statistically significantly lower than the national average. And *third*, it either does not have low reliability or has an average score more than one standard error below the 15th percentile. 42 C.F.R. § 422.166(a)(3)(i). Similar formulas follow for each of the five stars. Thus, RDS Testing is the means by which CMS turns every case-mix-adjusted CAHPS score into a measure-level rating.

Because of element (B) in the formulas above, CMS must determine the "national average CAHPS measure score" to perform RDS Testing. To do that, CMS gives greater weight to contracts with more enrollees and less weight to contracts with fewer enrollees. SUMMARY OF ANALYSES FOR REPORTING, MA & PDP CAHPS, https://ma-pdpcahps.org/globalassets/ma-pdp/scoring-and-star-ratings/2024/analysis_of_

reported_measures.pdf, last visited Aug. 8, 2025. Specifically, CMS "[c]alculate[s] the national mean (the weighted mean of all contract scores) for each measure, weighted by the survey-eligible enrollment assessed at the time of sample design." *Id.* ¶ 5.g. Because contracts differ in number of enrollees, CMS weights each contract by its number of enrollees and then takes the average of those values. In theory, this produces a similar result as taking the average of all the individual survey responses for a given CAHPS measure. It avoids giving disproportionate weight to scores in contracts with fewer enrollees.

Elevance claims it was arbitrary and capricious for CMS to use the "national weighted average" instead of the simple "national average." ECF No. 32 at 22. The RDS Testing regulation uses the phrase "national average CAHPS measure score." 42 C.F.R. § 422.166(a)(3)(i)–(v). Elevance contends weighting contracts by enrollees violates the regulation's requirement. On Elevance's interpretation, CMS must compare a contract's CAHPS score not to the average of all scores nationwide, but to the average of every contract's average score. In essence, the parties dispute whether the "national average" in the regulation is the average of scores at the enrollee level or the contract level. Elevance presents evidence that if CMS had used the average of all the contract-level scores, two of its contracts would have received higher Star Ratings. It argues that by using the weighted mean on top of the allegedly improper case-mix adjustments, CMS in effect puts a double penalty on providers.

1. Waiver

As a preliminary matter, CMS contends Elevance waived this argument. ECF No. 38 at 22. In its live pleading, Elevance does not mention that CMS weighted contracts by enrollment. Elevance pleads that "CMS violated its regulations and acted arbitrar[ily] and capriciously when it calculated CAHPS measures by first adjusting for the case-mix index, and then adjusting for the score's reliability and distance from the national mean, resulting in a double penalization." ECF No. 22 ¶ 67. The complaint does not mention the way in which CMS calculated the national mean. *See generally id.*

CMS cites *Johnson v. Thibodaux City* for the proposition that "the only claims that are relevant at the summary-judgment stage are those the plaintiff pleaded." ECF No. 38 at 22; *see* 887 F.3d 726 (5th Cir. 2018). Elevance counters that *Johnson* dealt with a plaintiff's ability to defeat, not obtain, summary judgment. But Elevance is presently seeking not only to obtain summary judgment but to avoid summary judgment in CMS's favor. So *Johnson* is, in fact, on point.

Elevance also argues that CMS has been on notice of its "weighted average" argument since it filed its summary-judgment brief. ECF No. 42 at 16. Under Rule 8, a pleader must make a "short and plain statement of the claim showing that the pleader is entitled to relief." FED. R. CIV. P. 8. While a plaintiff is not required to use "magic words" in its complaint, it must at least "give the defendant fair notice of what the claim is and the grounds upon which it rests." *Boudreaux v. La. State Bar Ass'n*, 3 F.4th 748, 756 (5th Cir. 2021); *Bell Atl. Corp. v. Twombly*, 550 U.S. 544, 555 (2007) (quoting *Conley v. Gibson*, 355 U.S. 41, 47 (1957)). Elevance's live pleading does not give CMS fair notice that it was challenging CMS's use of an enrollee-level national average. Nevertheless, because Elevance and CMS have fully briefed the issue, the Court will address it on the merits.

2. Regulatory Text

The merits of this claim can be resolved by the regulatory text alone. The regulation directs CMS to compare a contract's "average CAHPS measure score" to "the national average CAHPS measure score." 42 C.F.R. § 422.166(a)(3)(i)–(v). In each of paragraphs (i) through (v), subparagraph (B) uses "average CAHPS measure score" twice. First it refers to the average of that contract, and then it refers to the national average. In the first instance, the phrase clearly refers to the average of individual CAHPS survey response values. It means the average of the ratings that come back from all survey respondents for that measure. So when the regulation uses the phrase "average CAHPS measure score" a second time, it must be interpreted consistently. Thus, when preceded by the adjective "national," the phrase most naturally means the average of all survey response values throughout the country for that measure. This natural reading aligns much more with CMS's approach

than the approach proposed by Elevance. As CMS points out, to ignore the discrepancy in enrollment between contracts would be like taking the average of all fifty states' average heights in order to get the national average height, rather than adjusting for each state's population. In that scenario, a person from Wyoming would affect the average considerably more than a person from California would.

Again, Elevance never argues that using the enrollee-level average is per se arbitrary or capricious. It only contends that weighting contracts by enrollees violated the regulation. But CMS applied the regulation correctly. It calculated the national average CAHPS score, not the average contract CAHPS score. CMS is therefore entitled to summary judgment on this ground.

C. How to Round to the Nearest Half

CMS takes the average of all the measure-level ratings weighted by importance and produces that contract's raw score. It then rounds that number to the nearest half, using traditional rounding rules, to arrive at the final Star Rating. The regulations define "traditional rounding rules" as follows:

> Traditional rounding rules mean that the last digit in a value will be rounded. If rounding to a whole number, look at the digit in the first decimal place. If the digit in the first decimal place is 0, 1, 2, 3, or 4, then the value should be rounded down by deleting the digit in the first decimal place. If the digit in the first decimal place is 5 or greater, then the value should be rounded up by 1 and the digit in the first decimal place deleted.

42 C.F.R. § 422.162(a). This method is sometimes called "rounding half up."[3] Conceptually, this method asks whether an unrounded number is closer to the next-highest or the next-lowest whole number, with an upward tiebreaker. Thus, 1.3 rounded to the nearest whole is 1 because it falls between 1 and 2 and is closer to 1 than 2. Some numbers fall exactly between two increments. For example, 1.5 is just as close to 1 as it is to 2. A "tiebreaker" is necessary for such cases. The traditional

[3]Rounding Half Up, in ROUNDING, WIKIPEDIA, THE FREE ENCYCLOPEDIA, https://en.wikipedia.org/wiki/Rounding, last visited Aug. 8, 2025.

method breaks a tie by treating the unrounded number as if it were closer to the higher value; hence, "rounding up." Thus, 1.5 gets rounded up to 2. The tiebreaker is necessary only in such cases where there is no "nearest" whole number because the next-highest and next-lowest whole numbers are equidistant from the unrounded number.

The Star Rating regulations require CMS to round a contract's score not to the nearest whole number, but to the nearest half. This differs from the definition's whole-number example in two ways. *First*, the last digit in the value will be the first decimal place, not the ones place. *Second*, and more challenging, the score must be rounded not merely to the nearest number in the first decimal place, but to the nearest number with a 0 or 5 specifically in the first decimal place.

To round to the nearest half, then, the unrounded value will be rounded to whichever multiple of 0.5 it is closest to. For example, 1.3 rounded to the nearest half would be 1.5 instead of 1, as it is 0.3 away from 1, but only 0.2 away from 1.5. Values that fall exactly between increments of 0.5—that is, those ending in .25 or .75—get rounded up, per the tiebreaker. So 1.25 would become 1.5, even though it is equally close to 1 and 1.5.

Rounding to the nearest half can be validated by doubling the numbers, rounding to the nearest whole, and then dividing by two. For example, to round 1.3 to the nearest half, one could double it (2.6), round to the nearest whole (3), then divide by two (1.5). Rounding 1.25 would likewise be validated by doubling (2.5), rounding half up (3), and dividing by two (1.5). Thus, 1.25 rounded to the nearest half is 1.5.

Here, Elevance challenges the rounding on its contract H3655. Contract H3655 received a raw numeric score of 3.749565. The difference between 3.749565 and 3.5 is 0.249565. The difference between 4 and 3.749565 is 0.250435, which is slightly greater. 3.749565 is therefore closer to 3.5 than to 4. The result can be validated by the double-and-half method described above. Two times 3.749565 is 7.49913. The nearest whole number to 7.49913 is 7. Half of 7 is 3.5. CMS, therefore, correctly rounded the raw score of contract H3655 from 3.749565 to 3.5.

Applying the regulatory definition of "traditional rounding rules" confirms this result yet again. (The text of 42 C.F.R. § 422.162(a) appears in italic font.) *If rounding to a whole number, look at the digit in the first decimal place.* CMS must round to a half, which is in the first decimal place. Logically applying the rule, one would look to the digit in the second decimal place: 3.74̲9565. *If the digit in the first* (here, the second) *decimal place is 0, 1, 2, 3, or 4, then the value should be rounded down by deleting the digit in the first* (second) *decimal place. If the digit in the first* (second) *decimal place is 5 or greater, then the value should be rounded up by 1 and the digit in the first* (second) *decimal place deleted.* Here, the digit in the second decimal place is 4, so the 4 (and all following digits) are deleted—leaving only 3.7. Because the result must be a multiple of 0.5, the next-lowest value is confirmed to be 3.5.

Elevance's argument to the contrary is based on a fundamental misapplication of the rounding rules. Elevance argues as follows:

> [I]n order to most accurately determine the Star Rating for any contract, the midpoint between the half stars must be used—i.e., 0.25 and 0.75. . . . In other words, CMS's technical guidance explains that you should look to the second (or hundredth place) decimal to reach the midpoint. Therefore, by applying the regulatory definition of traditional rounding rules and basic math principles, you would look to the decimal to the immediate right of the second decimal (i.e., the third or thousandth place decimal) and round there. This would enable CMS to determine whether that score is closer to the lower or the higher half-star increment, consistent with 42 C.F.R. § 422.166(d)(2)(iv)'s clear requirement to determine star ratings in half increments. . . . Had CMS rounded to the hundredth (i.e., second) decimal, as would be appropriate when rounding to a half-star increment, H3655 would achieve a score of 3.75 and have been awarded 4 Stars.[4]

[4] The Court is at a loss to understand how rounding to the hundredth place would be "appropriate when rounding to a half-star increment." In fact, the final sentence quoted above grows more perplexing upon each reading. It is just as if one said, "Had Johnny divided by 100, as would be appropriate when dividing by 5 . . . ."

ECF No. 32 at 26–27. Elevance's proposed rounding method involves rounding twice. First, it holds that CMS should "look to the [third decimal] and round there." *Id.* at 26. That would determine whether a score would "reach the midpoint" or not. For contract H3655, this would mean rounding 3.749 to 3.75. Next, this proposed method would round 3.75 up to 4 because "traditional rounding rules" treat the midpoint by rounding it up. So Elevance contends 3.749565 should be rounded up to 4, even though it is closer to 3.5.

This explanation is gravely mistaken. Elevance acknowledges that the midpoint between 3.5 and 4 is 3.75. Crucially, the midpoint is not the number one is rounding to. Instead, is the **only** number that requires a special tiebreaker rule to round it up because it is equally close to the lower and higher one-half increments. Any number below the midpoint rounds down because it is closer to the next-lowest half, and any number above the midpoint rounds up because it is closer to the next-highest half. It is therefore perplexing when Elevance reasons that "by applying the regulatory definition of traditional rounding rules and basic math principles, you would look to the decimal to the immediate right of the second decimal . . . and round there." Elevance does not explain why it thinks one should look *past the midpoint*—that is, two digits to the right of the place to which you are rounding. Nor is it apparent what reason there could possibly be for doing so.[5] There is no need to look to the right of the midpoint; the midpoint (0.25 or 0.75) is itself already one decimal place to the right of the tenths place.

Following Elevance's proposed method would be akin to rounding 1.46 to the nearest whole number by identifying 1.5 as the midpoint and then looking to the number to the *right of the midpoint* and "rounding there." Thus, the 6 would raise the 1.46 to 1.5, which would then round up to 2. This approach would yield the absurd result of rounding 1.46 to 2, even though it is closer to 1. Likewise, Elevance advocates for the absurd result of rounding 3.749565 to 4, even though it is closer to 3.5.

[5]The Court suspects that Elevance has conflated "looking to" a digit in the traditional rounding rules definition with "rounding to" that decimal place.

As CMS points out, "there is no 'reaching the midpoint.' Either your score is above, below, or equal to the midpoint." ECF No. 38 at 30 (cleaned up). Elevance responds that the regulation "simply mandates the use of traditional rounding rules but does not foreclose CMS from applying its own policy to determine whether the rounded number is above, below, or equal to the midpoint of .75 or .25." ECF No. 42 at 28. One might as well apply one's "own policy" to determine whether 3 is greater than, less than, or equal to 4. Unfortunately, Elevance's argument appears to be tailored to reach the next Star Rating instead of anchored to the regulatory text.

If 3.749565 were in fact rounded at the second decimal (i.e., the hundredths place) as Elevance advocates for doing, the result would in fact be 3.75. But importantly, the regulation requires rounding not to the second decimal place but to the nearest half, which falls in the first decimal place. The number 3.75 is only relevant because it is the midpoint that rounds up to 4. Elevance does not contend that contract H3655's score is equal to or higher than 3.75. It contends instead that 3.749565 should round first to 3.75 and then to 4. If that is not double rounding, nothing is.

Elevance tries to wriggle out of the charge of double rounding. Elevance calls it a "red herring" and "irrelevant" and accuses CMS of doing the same thing. According to Elevance, CMS also rounds twice—once to the millionths place and once to the nearest half. ECF No. 42 at 26–28. As evidence, Elevance cites to a portion of the administrative record in which CMS states, "[t]he improvement measures, summary, and overall ratings are **calculated** with at least six digits of precision after the decimal whenever the data allow it." ECF No. 33 at App. 100 (emphasis added). CMS denies that it rounds to the sixth decimal place.

Elevance also points to CMS's Excel sheets as evidence that it rounds to the sixth decimal place. "As explained in Microsoft's support page, the FIXED function in Excel "[r]ounds a number to the specified number of decimals . . . and returns the result as text." *Id.* at 26 (quoting Microsoft Support, FIXED FUNCTION, https://support.microsoft.com/en-us/office/fixed-function-ffd5723c-324c-45e9-8b96-e41be2a8274a, last accessed Apr. 9, 2025). Elevance continues, "[I]f one were to remove the

FIXED function in the Score Card for H3655, the Overall Star Score numerical number would be 3.7495648235—which Defendants clearly rounded to 3.749565 when reporting it." ECF No. 42 at 27. Thus, Elevance argues, CMS cannot complain of Elevance's proposed double rounding because CMS double rounds itself.[6]

But Elevance confuses rounding the final Star Rating with rounding for display purposes. The purpose of rounding the final Star Rating is to simplify the measurement process. Otherwise, CMS would have to treat each contract according to a sliding scale rather than a set of increments. As a result, an enormous degree of administrative efficiency would be lost. Instead, by rounding to the nearest half star, CMS can place contracts into ascending tiers and quickly determine what benefit to award to certain providers from contract to contract. In contrast, rounding for display purposes (like Excel's FIXED function does) allows a number to fit in a cell on a spreadsheet. Because the raw score of a contract is the average of thirty or more weighted measure-level ratings, it will be an extremely precise number, capable of calculation to potentially infinite decimal places.

Tellingly, Elevance admits that CMS correctly rounds to the nearest whole number when rounding measure-level scores, even though CMS follows the same method for overall Star Ratings. ECF No. 42 at 25. Elevance argues that the purported difference between how CMS rounds measure scores and overall scores proves that CMS is acting arbitrarily and capriciously. Elevance quotes a selection from the Technical Notes about how CMS rounds the measure scores:

[6]This argument does not get off the ground in the first place. There is only one case in which such double-rounding would make a difference, even if CMS did perform it. Double rounding could make a difference if the raw average score was between 3.7499995 and 3.7499999. In that incredibly narrow range, rounding first to the nearest millionth place would take the score up to 3.750000, and rounding to the second place would take the score up to 4. In any other range, it makes no difference what happens to the right of the tenths place. But no double-rounding was performed in any event; CMS merely displayed the raw scores to six digits for convenience' sake.

> Measure scores are rounded using traditional rounding rules. These are standard "round to nearest" rules prior to cut point analysis. To obtain a value with the specified level of precision, the single digit following the level of precision will be rounded. If the digit to be rounded is 0, 1, 2, 3 or 4, the value is rounded down, with no adjustment to the preceding digit. If the digit to be rounded is 5, 6, 7, 8 or 9, the value is rounded up, and a value of one is added to the preceding digit. After rounding, all digits after the specified level of precision are removed. If rounding to a whole number, the digit to be rounded is in the first decimal place. If the digit in the first decimal place is below 5, then after rounding the whole number remains unchanged and fractional parts of the number are deleted. If the digit in the first decimal place is 5 or greater, then the whole number is rounded up by adding a value of 1 and fractional parts of the number are deleted.

ECF No. 42 at 25; *see* ECF No. 33 at App. 101. Elevance approves of this method. *Id.* But Elevance leaves out the very last sentence: "For example, a measure listed with a Data Display of 'Percentage with no decimal point' that has a value of 83.499999 rounds down to 83, while a value of 83.500000 rounds up to 84." ECF No. 33 at App. 101. But that is exactly what Elevance calls "rounding at the millionths place" when it comes to the overall Star Ratings. In fact, Elevance cites the same page of the Technical Notes as proof that CMS is "rounding to the millionth place" because it mentions 3.749999 and 3.750000. *Id.* at App. 101. What is sauce for the goose is sauce for the gander. Elevance recognizes CMS's rounding method for measure scores is correct. But when CMS uses the exact same approach for overall Star Ratings, Elevance accuses it of violating the law.

One more point illustrates why this argument fails. Consider the result if Elevance had its way. (The Court agrees with and adopts the demonstrative charts in CMS's brief.) The chart on the left records the raw scores that get converted to each final summary or overall Star Rating under CMS's approach. The chart on the right shows the same thing under Elevance's proposed approach. CMS is correct that Elevance's proposed method is essentially gerrymandered to give it the result it wants.

Codified Star Ratings Methodology
(Traditional Rounding Rules)

Raw Summary / Overall Score	Final Summary / Overall Rating
≥ 0.000000 and < 0.250000	0
≥ 0.250000 and < 0.750000	0.5
≥ 0.750000 and < 1.250000	1.0
≥ 1.250000 and < 1.750000	1.5
≥ 1.750000 and < 2.250000	2.0
≥ 2.250000 and < 2.750000	2.5
≥ 2.750000 and < 3.250000	3.0
≥ 3.250000 and < 3.750000	3.5
≥ 3.750000 and < 4.250000	4.0
≥ 4.250000 and < 4.750000	4.5
≥ 4.750000 and ≤ 5.000000	5.0

Elevance Double Rounding

Raw Summary / Overall Score	Final Summary / Overall Rating
≥ 0.000000 and < 0.245000	0
≥ 0.245000 and < 0.745000	0.5
≥ 0.745000 and < 1.245000	1.0
≥ 1.245000 and < 1.745000	1.5
≥ 1.745000 and < 2.245000	2.0
≥ 2.245000 and < 2.745000	2.5
≥ 2.745000 and < 3.245000	3.0
≥ 3.245000 and < 3.745000	3.5
≥ 3.745000 and < 4.245000	4.0
≥ 4.245000 and < 4.745000	4.5
≥ 4.745000 and ≤ 5.000000	5.0

In conclusion, CMS was correct to round 3.749565 to 3.5. That is because, simply put, 3.749565 is closer to 3.5 than to 4. CMS is therefore entitled to summary judgment on Elevance's arbitrary-and-capricious claim as to the alleged rounding error.

Elevance brings up two more arguments built on the premise that CMS rounded to the nearest millionth place. Those arguments will necessarily fail because, as discussed above, CMS does not round to the nearest millionth; it calculates far beyond the millionth place, displays the number in the Technical Notes, and then rounds to the nearest half. But for the sake of thoroughness, the Court will address each argument.

D. Effect of Statistical Variance on Star Ratings

Elevance argues that rounding to the nearest millionth is arbitrary and capricious because it ignores the inherent imprecision in calculating non-CAHPS and CAHPS measures.

Elevance argues that mean resampling of non-CAHPS measures introduces too much statistical imprecision to justify rounding to the nearest millionth. When scoring non-CAHPS measures, CMS uses cut-point analysis with mean resampling. *See supra*, Statutory and Regulatory Background. CMS's process aims to place the "cut points," (i.e., the gaps between each increment from one to five) in such a way that the all the raw scores within the same cluster are as close as possible to each other and as far as possible from those in the next highest or lowest cluster. 42 C.F.R. § 422.166(a)(2)(i). Mean resampling

aims to make the cut points less sensitive to statistical outliers. *See* ECF No. 33 at App. 33 n.18. It works by randomly sorting the values of a given measure across all MAOs into ten equally sized groups. The average is taken ten times, leaving one of the ten groups out each time.

Elevance provides the declaration of proposed expert Dr. Paul Diver, also of Berkeley Research Group, who holds a Ph.D. in Statistics. ECF No. 33 at App. 27 ¶ 12. Dr. Diver explains that mean resampling will produce different cut points depending on the groups into which the contracts' raw scores get sorted at the beginning of the process. ECF No. 33 at App. 27 ¶ 20. The scores will be sorted differently depending on what seed number is used. *Id.* The seed therefore introduces an element of randomness to the cut points. Whether a contract falls on one side of a cut point or another determines its measure star rating for that measure. Thus, random chance can make a difference as to a contract's measure-level rating if it is close enough to the cut point. ECF No. 33 at App. 32 ¶ 22–23. This does not affect the measure-level scores only. Because the ultimate Star Rating is a function of the measure-level scores, the use of a different seed in the mean resampling process can result in a variation in the final score. ECF No. 33 at App. 35–36 ¶ 30. Dr. Diver gives an example in which a contract's overall score may vary from 3.749565 to 3.769173 based on a change in the seed alone. *Id.*

Importantly, Elevance does not argue it is arbitrary and capricious for CMS's calculation to have any degree of statistical variation. Instead, it argues that *given* the degree of variance, it is arbitrary and capricious for CMS to round overall scores to the sixth decimal place, or nearest millionth. ECF No. 32 at 28.[7] Elevance contends that by rounding to the nearest millionth, CMS "attempt[s] to create a false sense of precision in the Final Summary Score." *Id.* Instead, Elevance contends CMS should have accounted for the inherent statistical variance by rounding the scores to the second decimal place, or nearest hundredth. Those arguments fail for two reasons.

[7]"Calculating and rounding a contract's Final Summary Score to the sixth decimal is arbitrary and capricious because it ignores the imprecision of the methodology used by CMS to determine a plan's Final Summary Score."

First, CMS rounds to the nearest half—not to the nearest millionth—resolving any potential "false sense of precision." Elevance complains that the sixth decimal place is too precise to round to in light of statistical variance and says the rounding should be less precise than 0.01. It therefore proposes the second decimal place. But Elevance does not contest that the final score is rounded to the nearest half-star increment, which is even less precise than 0.01. So while it is clear what Elevance specifically would gain from rounding to the second decimal place (i.e., a half star), it is unclear how its proposed rounding scheme would be any better at accounting for the alleged statistical variance.

Second, statistical variance would still make a difference in Star Ratings even on Elevance's proposed approach. Suppose CMS did as Elevance proposes and rounded each score to the second decimal before rounding to the nearest half-star increment. In such a case, the statistical variance caused by the use of different seeds would still exist. In fact, it might appear not to exist, hidden in the process of rounding. For example, a single contract may get two different raw scores of 3.745000000 and 3.744999999, respectively, based on a change in the seed number. That is a difference of only 0.000000001, or one billionth. Yet, if CMS rounded those numbers to the hundredths place, as Elevance proposes, those scores would become 3.75 and 3.74, respectively.[8] Once Elevance's proposed second rounding was performed, a difference of a half star would still exist in the final Star Ratings, even though the variation was only one billionth of a point. No matter what digit one rounds to, imprecision is inevitable.

Elevance's argument based on sampling error in the CAHPS survey data is yet another verse in the same song. Again, Elevance's problem is not that CAHPS survey data is prone to error; it is that "CAHPS survey data is not fit for making Star Ratings decisions that rest upon differences at the *millionth* decimal point." ECF No. 32 at 33. Elevance produces the declaration of another expert, Dr. Paul L. Lavarkas, Ph.D., "a world-renowned survey research expert." ECF No. 32 at 30. Dr.

[8]Rounding 3.7**4**4999999 to the second decimal place would require one to look to the next digit, which here is less than 5, and thus round down.

Lavarkas attacks the CAHPS survey data based on three main claims. First, CMS failed to investigate and account for nonresponse bias. *Id.* Second, CMS failed to account for inherent sampling error. *Id.* at 31. Third, CMS may not have eliminated "problematic" survey questions that potentially contributed error and imprecision in the reliability of respondents' answers. *Id.* at 33.

Once again, even assuming those allegations are true, Elevance's argument fails because it rests on a faulty assumption. CMS does not round to the sixth decimal place. And if CMS performed the double rounding Elevance proposes, it would not change the fact that some contracts would receive different scores based on slight differences at the margins. Elevance is clear that it does not criticizing the CAHPS survey methodology per se, but merely the decision to "calculat[e] Star Ratings to such an extreme" in light of the alleged survey flaws. *Id.*[9]

On a final note, Dr. Diver says the average statistical uncertainty in CMS's method is 0.01. ECF No. 33 at App. 26 ¶ 8.i. He opines that contract H3655's score (3.749565) falls short of 3.75 by only 0.000435, which he says is a difference "orders of magnitude smaller than the average statistical uncertainty due to random chance inherent in CMS's methodology." *Id.* at App. 26 ¶ 8.v. But that is not the right difference to focus on. Instead, the right comparison to make is between 3.749565 and 3.5 (0.249565) which is less than the difference would be if CMS had rounded contract H3655's score up to 4 (i.e., 0.250435). The difference due to rounding to the nearest half star—0.249565—is much greater than the statistical uncertainty of 0.01 that exists according to Dr. Diver.

---

[9]Elevance's cited case is inapposite. *See Friends of the Boundary Waters Wilderness v. Bosworth*, 437 F.3d 815, 826–27 (8th Cir. 2006). There, the agency acted arbitrarily and capriciously when it calculated the motorboat use rate. *Id.* at 824. The agency conducted a survey of 13 people asking about how frequently they used their motorboats in 1978, twenty years before the survey was taken. *Id.* at 825. Only five people responded to that question. *Id.* The court's decision did not have to do with the agency's choice to round a value to a certain degree of precision in light of poor survey quality. It was merely about the poor quality of the survey. For that reason, *Friends of the Boundary Waters Wilderness* is of no importance here.

Because CMS rounded to the nearest half star, and not the nearest millionth, Elevance's argument fails. CMS is entitled to summary judgment on this claim.

E. The Jenny Seed

Elevance's final argument is a rehash of its argument in the prior section about the variation caused by mean resampling. Conceptually, it belongs more in the prior section than in its own. But because Elevance places this argument in its own section in its brief, the Court will analyze it separately.

Elevance complains that because CMS uses the "Jenny Seed" to kickstart the mean resampling process every year, the assortment is not truly random as required. Again, Elevance claims the repeated use of the Jenny Seed violates the regulations only in light of CMS's purported rounding to the millionth place.[10] But CMS does not round to the millionth place; it rounds to the nearest half. *See supra Section C.* So Elevance's argument fails for that reason alone—but not for that reason only.

To perform mean resampling, the regulation requires that CMS must first "randomly separate[ ]" measure-specific scores "into 10 equal-sized groups." 42 C.F.R. § 422.162(a). CMS performs the random sorting process by starting with a "seed value." ECF No. 33 at App. 185. According to CMS's Technical Notes, "[t]he *seed=8675309* option [in the computer program] specifies the seed value that controls the starting point of the random sequence of numbers and allows for future replication of the randomization process." *Id.*

Elevance argues that by using the same seed year after year, CMS "defied the regulatory requirement that mean resampling be random." ECF No. 32 at 35. According to Elevance's expert, Dr. Diver:

---

[10]"Rounding to the Millionth Decimal Also Violates CMS's Own Regulation By Calculating Non-CAHPS Cut Points By Using the Same Seed of 8-6-7-5-3-0-9 Year Over Year." ECF No. 32 at 34. "This fundamental error in CMS's methodology [using the Jenny Seed every year] further illustrates why CMS's decision to round Star Ratings to the *millionth* decimal is arbitrary and capricious." *Id.* at 35.

> With that [the repeated use of the same seed] in mind, the composition of the plan groupings would therefore be predictable and not randomly different from one year to the next without changes to other external factors such as the population of plans considered or the initial ordering (e.g., alphanumerically). In other words, if the list of considered plans remained the same from one year to the next, and the same seed is used each year, the plan grouping in CMS's clustering methodology would be effectively pre-determined year over year.

ECF No. 33 at App. 54.

But even taking Dr. Diver's testimony at face value, it does not follow that CMS's method violates the randomness requirement. First, as CMS points out, the seed would not generate the same random groupings if the list of MAO contracts changed. And according to CMS's fact witness, Dr. Elizabeth Goldstein,[11] "[t]he set of contracts subject to mean resampling differs from year to year." ECF No. 51 ¶ 32. So even if the Jenny Seed were capable of producing the same ten groups in theory, it did not do so in reality.

But the Jenny Seed does not necessarily depend on a different lineup of contracts every year. For instance, suppose that for two consecutive years, the list of MAO contracts remained perfectly unchanged. Assuming they were listed the same way before being sorted (for example, alphanumerically), the ten groups would be the same in both years. It does not follow that the measure scores in the second year's analysis would not be random. Elevance makes the jump from the language in Dr. Diver's declaration ("not *randomly different*") to "*non-random*." ECF No. 33 App. 54 ¶ 88; ECF No. 32 at 35. The Court expresses no opinion as to whether the sorting could still be considered "random" if it resulted in the same ten groups twice. The Court merely notes that Elevance assumes, without proof, that it would not be.

[11] Director, Division of Consumer Assessment and Plan Performance, Medicare Drug Benefit and C & D Data Group, Center for Medicare, CMS. ECF No. 51 ¶ 1. Dr. Goldstein is a fact witness, not an expert witness. *Id.*

Additionally, it is difficult to imagine how the use of the same seed year after year could have harmed Elevance. Dr. Diver states that the use of another seed could have resulted in a higher final Star Rating for contract H3655. That may be. But Elevance does not argue that the Jenny Seed was the *only* seed that would have resulted in a Star Rating of 3.5. Nor does it argue that there is some legal reason why CMS should have used a specific seed that would have pushed it over the line to 4 stars. Even if CMS were required to use a different seed every year, the Court could not conclude that its failure to do so harmed Elevance. Suppose CMS had used the Jenny Seed for the first time when calculating the 2025 Star Ratings; before that, a different seed was used. Elevance's position would be no different, even though the alleged violation of law would not have taken place.

For those reasons, Elevance's argument based on CMS's use of the "Jenny Seed" fails. Again, though, even if Elevance's argument were sound, it would be rejected because it assumes that CMS rounds to the millionth place, which is not the case. In any event, CMS is entitled to summary judgment on Elevance's arbitrary-and-capricious claim as to the use of the Jenny Seed.

* * *

In the Star Rating system, it is virtually guaranteed that there will be contracts that fall a hair's breadth short of the next-highest rating. In the 2025 Star Ratings, that appears to have happened to one or two of Elevance's contracts. Understandably, Elevance would have preferred for those contracts to reach the next half-star tier. But the fact that they fell short does not give rise to a claim for relief under federal law.

This case has only scratched the surface of the complex process by which CMS collects data and produces Star Ratings. That process involves advanced knowledge of data collection, statistics, and mathematics. In other words, it is not one which a federal court is well suited to second guess. Absent any arbitrary and capricious conduct by CMS, this Court is in no position to question the outcome of the Star Rating system. Here, Elevance has not shown any evidence that CMS acted arbitrarily or capriciously.

CONCLUSION

Therefore, the Court **GRANTS** CMS's Motion for Summary Judgment (ECF No. 37) and **DENIES** Elevance's Motion for Summary Judgment (ECF No. 31). Each party shall bear its own costs.

SO ORDERED on this **18th day of August 2025.**

Mark T. Pittman
UNITED STATES DISTRICT JUDGE